

Detecting shifts in metropolitan structure: a spatial network perspective

Journal Title
XX(X):1–17
©The Author(s)
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/

SAGE

Abstract

The spatial distribution of metropolitan areas creates a complex system of competing and cooperating economic units. Understanding the size, scale and functional structure of such systems is an important topic in regional science. This paper uses commuter flow data over a forty year period to analyse the changing structure of the Greater Manchester metropolitan area. We apply a combination of complex network analysis, residual network analysis and spatial network visualisation to detect Greater Manchester’s polycentric structure and identify intra-regional communities. This method is able to identify economic geographies in a highly complex and interdependent commuter network. We comment on the role of administrative boundaries in shaping metropolitan regions and discuss the potential of our work to inform debates on regional governance geographies and local government planning practices.

Keywords

Community detection, network analysis, metropolitan areas, residual networks, Greater Manchester.

Introduction

Metropolitan areas have emerged as important economic and political spaces. They often take a polycentric form, reflecting the urban agglomeration processes that shape contemporary urban development (Fang and Yu 2017). Recent work in urban theory has concluded that the complexities of such urban systems are best understood through the interactions that shape them (Batty 2013; Pflieger and Rozenblat 2010). Adopting this theoretical stance encourages urban systems to be studied using a complex network framework that focuses on the set of relationships that define a system. Whilst topological relationships, including political or social interactions, influence the form and extent of the interplay between proximate areas, it is also important to acknowledge the importance of geographic

Corresponding author:

Email:

space in facilitating economic and social exchanges (Berdegué et al. 2019). We apply a combination of complex network analysis, spatial interaction models, and geographic network visualisation to understand the functional structure of metropolitan areas and how they change over time. Our work is motivated by the following research questions:

- To what extent, and in what ways, can the functional structure of metropolitan areas be discerned from the patterns of human activity revealed by commuter flows?
- How has metropolitan structure in the UK changed over the last four decades?
- How can spatial network analysis inform a debate on metropolitan policy making and the configuration of regional governance boundaries?

We address these research questions through empirical analysis of commuter flows in Greater Manchester (Figure 1). Commuter flow data can be extracted and interpreted using network analysis, providing a way of delimiting meaningful labour market areas that can inform economic development policy and land-use planning (Hincks et al. 2018).

Greater Manchester was chosen as the case study area due to its claimed status as an exemplar of contemporary metropolitan governance (Haughton et al. 2016; Harding et al. 2010). Following three decades of collaborative governance under the Association of Greater Manchester Authorities (AGMA), Greater Manchester became the first statutory combined authority to be created in England. As such, subsequent devolution deals agreed between Greater Manchester Combined Authority (GMCA) and central government mean that city-regional governance is more thoroughly developed than in other parts of England (Kenealy 2016).

Our approach builds on methods developed in previous work (Nelson and Rae 2016; Zhong et al. 2014). We argue that the nature of intra-regional patterns of connectivity are more interdependent than regional or national patterns which were the focus of previous studies. As such, space-independent community detection methods struggle to depict the underlying functional structure that is evident in Greater Manchester's distribution of economic nodes. To account for these complexities, we extend previous work by using spatial interaction models to incorporate the geographic nature of urban networks into our analysis. Specifically, we use gravity models to refine the commuter network into a network subgraph, referred to as a residual network, that is more readily interpretable by community detection methods. Having detected major shifts in functional structure over time, our analysis culminates with a model of Greater Manchester that delineates the metropolitan area into functionally coherent communities. We then make comparisons between the network-based communities and Greater Manchester's existing administrative boundary structure.

The rest of the paper is organised as follows. In the next section, we review previous research on spatial network analysis and regional delineation and argue that this work extends our understanding of functional economic geographies in both methodological and empirical terms. In Section 3 we introduce the data used, explain the network extraction method employed, and provide a technical description of the network tools used in the analysis. We present results in Section 4 before reflecting on our approach and its limitations in Section 5. We conclude in Section 6, highlighting areas for future research.



Figure 1. Map of local authorities comprising Greater Manchester. Manchester City Centre ward highlighted to show the focal point of the city region.

Background

Spatial network analysis

Urban morphology, as Batty defines it, is the “study of form and structure in cities that focuses on the dynamics of change and the rules underlying these dynamics” (Batty 2013, pg.179). This encompasses the set of relationships arising from the interactions of people, freight, materials, and information (Rodrigue 2016), with the spatial structure of modern cities shaped in part by the increasing complexity of these interactions.

A network approach provides a framework for the study of such interactions and has been applied extensively to urban networks. Urban street networks have been particularly well studied, with early research from Jiang and Claramunt (2004) and Porta et al. (2006) analysing the network properties of urban street networks in a range of cities. Boeing (2019) later examined the orientation of street networks to measure the spatial order of cities.

The emergence of urban interaction data has been used to study the patterns of connectivity and human interactions that shape modern cities. For example, Roth et al. (2011) used Oyster card data to study the polycentric configuration of London, while Thiemann et al. (2010) used a dataset obtained from the online bill-tracking game wheresgeorge.com to study spatiotemporal phenomena generated by human activity.

To generalise traditional network approaches to account for the spatial characteristics of urban networks, recent research has used flow data or distance data to produce models that are more reflective of the spatial nature of urban networks. For example, Cheng et al. (2015) used network centrality measures that incorporate flow volume to understand Singapore’s mass rapid train network beyond its topological structure. Zhong et al. (2014) applied spatial network analysis to smart card data to identify hubs, centres, and borders in Singapore.

Despite being granular, novel interaction data is limited by its short history, with the effect that changes in urban structure over time cannot fully be understood. In this respect, we make two contributions to previous works. First, we use census interaction data to capture the long-term shifts in spatial structure that are characteristic of urban areas. And second, we employ gravity models, which have not been used in this context before, to add a spatial dimension to the network analysis.

Regional delineation

Regional delineation is an important topic for multiple disciplines, featuring in previous research by geographers, economists and urban planners (Sng et al. 2018). This makes for a diverse field of study, with many different approaches to defining regions and delineating boundaries. However, due to the complexity and heterogeneity of features that influence the configuration of boundaries (Jiang 2018), the development and operationalisation of a practical framework for regional delimitation is not straightforward. The complexity of the problem has long been recognised, with Hartshorne (1939) warning that the face of the earth is the “antithesis of a mosaic” and Nelson and Rae (2016) later commenting on the impossibility of defining stable territorial regions whose boundaries are uncontested. Any attempt to employ an algorithmic approach to the definition of regional political space therefore faces significant challenges.

Traditional approaches to regional delineation tended to employ a core-based approach, where an emphasis is placed on defining an urban core and identifying its connections to the periphery (He et al. 2019). Approaches of this nature are generally less interested in the patterns of connectivity and spatial agglomeration processes that shape the development of polycentric regions. Plane (1981) pioneered the use of a flow-based approach to show that suburbs were less dependent on central cities than it was previously supposed. Later, Pflieger and Rozenblat (2010) asserted the importance of networks in creating interdependencies between urban spaces. They proposed the view of “The city as a network of networks”, allowing urban systems to be studied using a complex network framework. Capitalising on this network perspective of regions, Nelson and Rae (2016) and Hamilton and Rae (2018) used a network partitioning approach to delineate regions in the U.S. and Scotland, respectively.

Other studies have emerged that do not use a network-based approach. These empirical approaches have been guided by the nature of the data and the scale of the analysis. Sng et al. (2018) opted for a geographic approach, using Geographic Information Systems and digital elevation data to create hypothetical provinces in China. Berdegué et al. (2019) used nighttime satellite imagery to identify the boundaries of conurbations. Other types of data that have been explored include sub-regional housing and labour markets (Hincks 2012), smart card data (Zhong et al. 2014) and mobile phone data (Louail et al. 2014).

Whilst many of these efforts have highlighted functional urban areas that differ greatly from existing administrative boundary structures, in the UK there has been a general reluctance to engage in wide-ranging local government reorganisation to create more logical geographies (Demazière and Sykes 2020). There have however, been experiments and reforms to regional governance, initially on a voluntary basis but more recently involving legislation, including the creation of GMCA in 2011. GMCA brings together the ten districts that previously constituted the Greater Manchester Council (GMC) and assumes responsibility for the region’s transportation, economic development, strategic land-use planning, and policing. GMCA boundaries were the result of political compromise, rather than any attempt to delimit boundaries in a more systematic way. The city-region’s political leaders discussed multiple

configurations, but eventually reverted to the ten district geography of the former GMC in order to sidestep potential sensitivities among neighbouring local authorities, fearful of Mancunian expansionism (Deas 2014). In this regard, it is unclear how well Greater Manchester’s administrative boundaries capture the functional reality of the region. To address this, we attempt to delineate Greater Manchester into functionally coherent communities that are reflective of its polycentric economic geography, augmenting the existing literature on regional delineation by utilising residual networks in this process.

Materials and methods

Origin-destination data

In this paper, we use data on the commuting flows of workers in the Greater Manchester metropolitan area. In an attempt to understand the functional changes of Greater Manchester across time, our analysis includes data from the four most recent UK census years (1981, 1991, 2001 and 2011). These data were obtained using the Web-based Interface to Census Interaction Data (Stillwell 2006). We consider travel to work data only, limited to persons aged 16 or over who were in employment in the week before each respective census year. The specific census question we are interested in concerns a person’s ‘usual residence and place of work’.

We perform our analysis at ward-level, where a ward is a subdivision of a local authority area and has an average population size of 5,500 people. This geographical unit is comparable to those used in previous studies, such as US Census Tracts (Nelson and Rae 2016) and Scottish Intermediate Zones (Hamilton and Rae 2018). Ward boundaries and nomenclature are consistent across the first three time periods but have minor changes for 2011. As we are interested in the structural properties of the commuter network, these changes have a negligible effect on our analysis and we do not perform a geographical lookup or aggregation to construct consistent ward boundaries. We exclude flows that originate or terminate outside of the Greater Manchester area.

Network extraction

We use the ward-level origin-destination matrix to construct a commuter network. The network consists of a set \mathcal{V} of N nodes (wards) and a set \mathcal{E} of M edges (an edge between two nodes will exist if there is at least one commute between an origin-destination ward pairing), and will be defined as a graph $G = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{n_1, n_2, \dots, n_N\}$. Every directed edge (i, j) in \mathcal{E} will have a weight $\omega_{i,j}$, equal to the number of commutes from ward i to ward j . N , M and $\omega_{i,j}$, $\forall i, j, i \neq j$ are determined separately by the data for each census year.

Table 1 shows basic network statistics for each year of data. There is a significant increase in network edges over time. This is attributable to an increase in commuters as a product of the population growth across the area. An increase in edges has impacted on the other statistics presented, with network snapshots showing an increase in average node degree and network density. Similarly, reciprocity and transitivity have also increased. These measures indicate the probability of two wards being mutually linked and the probability for the network to have adjacent wards interconnected, respectively. These increases are intrinsically linked with the increase in network size and do not affect our main methods performed. Opting for ward-level as the spatial unit of analysis results in a sizeable network of 214 nodes for 1981, 1991 and 2001, and 215 nodes for 2011. This consistency allows for comparative analysis to be performed.

Network Properties	1981	1991	2001	2011
Nodes	214	214	214	215
Edges	13297	13968	26345	35840
Average degree	62.14	65.27	123.12	166.70
Density	0.292	0.306	0.58	0.78
Reciprocity	0.508	0.485	0.677	0.83
Transitivity	0.575	0.583	0.811	0.93

Table 1. Basic network statistics for each year of data.

Complex network analysis

We analyse Greater Manchester’s commuter flows using network analysis. Our network extraction and centrality analysis is written in Python and uses graph-tool functions implemented in the Networkx library (Hagberg et al. 2008). We perform two methods of community detection, COMBO and InfoMap, using implementations in C++ (Sobolevsky et al. 2014) and R (Csardi and Nepusz 2006). Our analyses focus on identifying network structure and important nodes and communities. We are particularly interested in how these network characteristics have changed across the four time periods.

Node-level importance. Identifying the most important nodes in a system is a major element of network characterisation (Newman 2010). We implement a selection of centrality measures and algorithms that capture various aspects of what makes a node important, such as a node’s position in the network and the importance of its connections. For many spatial networks, these characteristics are closely linked to the morphology of the network (Barthélemy 2011). Such analysis is meaningful in the context of commuter flows as the structure of the network and the position which a node takes within the global topology has implications for policy design, patterns of mobility, and the interactions of goods and ideas (Berdegue et al. 2019). As we use well established centrality measures and algorithms in our analysis (node strength, eigenvector centrality, PageRank centrality and HITS algorithm), we present these in Supplementary 1.

Community detection. Community detection refers to the procedure of identifying groups of interacting nodes in a network depending on their structural properties. One of its aims is to identify groups of nodes that are more densely connected internally than they are with the rest of the network (Newman 2010). Identifying network communities in this way has implications for metropolitan areas. In our analysis, communities are theoretical groupings that offer insights into the functional structures that constitute metropolitan areas and the delineations of administrative boundaries. We implement two community detection algorithms that were chosen for their ability to account for the directed and weighted edges of the commuter network.

1. COMBO algorithm is a universal optimization technique for community detection. It is computationally expensive but has been shown to effectively recover community structure in a variety of networks, including commuter networks at large scales (Nelson and Rae 2016; Hamilton and Rae 2018). COMBO is a modularity maximisation technique, where modularity is a measure of the strength of division of a network into communities (Newman 2010). In their paper, Sobolevsky et al. (2014) attribute the success of the algorithm to the method in which vertices are recombined between communities.

2. Infomap is an algorithm that utilises random walks to detect communities. It is based on the premise that once a random walker enters a community structure it tends to stay within that structure for a long time, with movements between communities occurring relatively rarely (Esquivel and Rosvall 2011). It differs from COMBO algorithm in that it attends to patterns of flow on the network and the modularity maximisation approach, utilised by COMBO, does not. As such, the algorithms are prone to capture different elements of community structure (for further technical details of the InfoMap method, see (Rosvall et al. 2009)).

Residual network. Residual networks are network subgraphs. They are obtained by retaining only those edges in which their observed edge weight exceeds their expected edge weight by a given threshold. The expected edge weight can be calculated in a variety of ways and is dependent on the nature of the data. In the case of commuter networks, we calculate the expected edge weight using a gravity model. We formally define the expected commuter flow between two wards, F_g , as

$$F_g = G \frac{m_1 m_2}{r^2} \quad (1)$$

where G is a constant multiplier, m_1 and m_2 are the in-strength and out-strength of ward 1 and ward 2, respectively, and r is the euclidean distance between the wards. The multiplier G is calculated from the data. We take each directed edge and, using the length of the edge and the in-strength and out-strength of its adjoining nodes, calculate the constant needed in order to produce the observed edge weight, storing the values in a vector. From this vector of expected edge weights, we take the geometric mean, which is appropriate for heavy-tailed data, as our constant multiplier, G .

With the constituent elements calculated, we can now use Equation 1 to predict the expected commuter flow from one ward to another based solely on their respective node-strengths and distance from each other. A positive residual indicates that an edge has a greater flow than expected. Taking the residuals, we can produce *residual networks* that consist only of edges that exceed their expected flows at different thresholds. For example, a 5% residual network is a subgraph of the original networks that retains the top 5% of highest valued residual edges, in absolute terms, in the network. The value of the threshold is application dependent and is determined through visual interpretations of network plots, with the aim to strike a balance between network refinement and the preservation of community structure.

The residual network can be understood as representing the most important and interesting edges in the network. More significantly for metropolitan areas, however, is the structure of the residual network. The residual structure refines the global topology into clusters of edges that identify economic clusters and functional communities.

Visualisation. Spatial network visualisation and mapping in this paper relates to three main types of plots:

1. In the first set of maps, centrality values of network nodes are projected on to their respective ward-level polygons. This mapping shows how the geographical position of a node in Greater Manchester relates to its importance and function from a network perspective. The visualisations also describe how some elements of network topology relate to urban morphology such as infrastructure.
2. Visualising residual networks has the advantage of revealing the underlying structure of spatial networks when correcting for distance. Each edge in the visualisations represents a commuter

flow between two wards that were higher than expected given the gravity model. We show how these visualisations can be used to reveal functional structures at different thresholds and also how residual network plots can be used to inform the community detection process.

3. In regards to making a visual comparison of the existing administrative boundaries against potential new regional delineations, we adopt a similar approach to that of [Sobolevsky et al. \(2014\)](#). That is, we overlay an outline of the existing administrative boundaries (shown in black) on top of colour coordinated communities. The colour coordination facilitates comparisons to be made across years and across methods, revealing the changing structure and emergence or disappearance of communities.

Results

Centrality and clustering: hubs, centres and shifts in structure

To explore how the topological structure of Greater Manchester relates to the geographical space in which it is embedded, we project measures of node-level importance onto their respective ward polygons. [Figures 2-5](#) are the maps of the eigenvector, PageRank, authority, and clustering coefficient scores for each time period. The emerging structure is one of polycentricity, with clear shifts in structure and importance for many areas of the region.

Analysing the results more closely, we see that [Figures 2, 3 and 4](#) capture similar network characteristics. As would be expected, geographically central wards constitute a nexus of economic activity, with peripheral wards, for the most part, being less important. This pattern reflects the concentration of service sector organisations in the city centre. There are, however, wards located elsewhere in the region that are also of markedly higher importance than their neighbours. These centres stand as secondary local attractors of commuters to their respective poly-centres and act as important economic units within the wider region. Observing only [Figures 2, 3 and 4](#), there is inconclusive evidence to support major structural changes in Greater Manchester's urban morphology over time; individual wards wax and wane in importance, but the global structure remains relatively consistent.

Unlike the previous centrality measures, the distribution of clustering coefficient values in [Figure 5](#) suggests significant changes in connectivity patterns over time. The local clustering coefficient quantifies how close a ward's neighbours are to forming a clique. In reality, then, a highly clustered area is indicative of commuters taking up employment opportunities in neighbouring wards. In 1981, 1991, and 2001, Wigan is the most clustered area. Given that, in terms of physical distance, Wigan is furthest from the centre of Greater Manchester, it is conceivable that residents of Wigan are more likely to seek economic opportunities in a more localised area. Over time, however, and culminating in 2011, commuting patterns involving Wigan become markedly more integrated, and the central Manchester area emerges as the most highly clustered area.

In the next section, we discuss the results of our main metropolitan structure analysis. We expect COMBO and InfoMap to detect community structures that reflect the polycentricity identified in this section.

Community detection for intra-regional delineation

Administrative boundaries often define regions on the basis of historical patterns of identity and inherited political subdivisions. Although this may hinder the ability to govern effectively, updating territorial

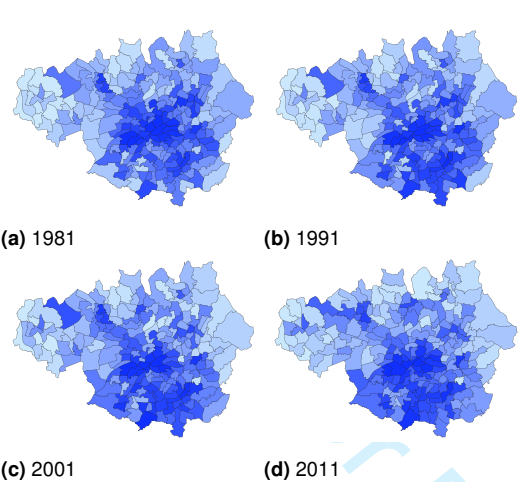


Figure 2. Eigenvector centrality scores

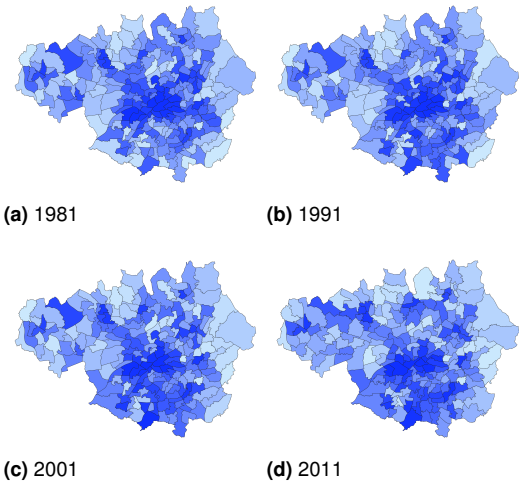


Figure 3. PageRank centrality scores

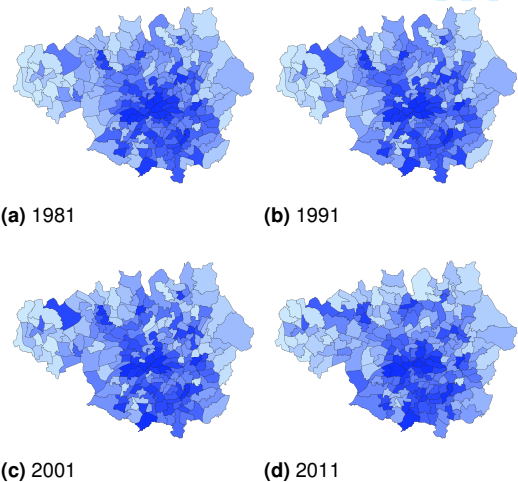


Figure 4. Authority scores

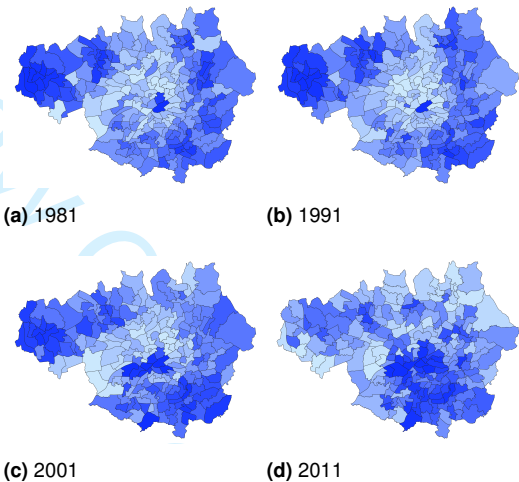


Figure 5. Clustering coefficient scores



boundaries may not be a political priority, for two principal reasons: the practical and political difficulty of restructuring administrative boundaries and the absence of a robust framework to determine functional boundaries to which different actors can subscribe. In this section, we address this latter issue by employing community detection algorithms to delineate intra-regional borders based on community

structure. We track the changes in community structure over time and highlight discrepancies with Greater Manchester's current administrative boundaries.

Figure 6 maps COMBO communities onto the Greater Manchester region, with the existing local authority structure included for comparison. The results of COMBO for 1981 divide Greater Manchester into seven communities. The COMBO communities for Wigan, Tameside, and Stockport are identical or very similar to the existing administrative structure, but the remaining four COMBO communities span multiple existing local authorities. The results for 1991 and 2001 are similar, with the exception of the introduction of a new COMBO community that maps onto the Bury area. For 2011, however, the community structure changes significantly. The COMBO communities along the northern boundary remain stable but two southern communities disappear when compared to 2001. The disappearance of these two communities is compensated by the growth of two existing communities that previously occupied Manchester and Trafford (to which Salford is added in 2011), and Stockport (to which Tameside is added). The notion that the commuting patterns become more integrated across this central and southern area is indicated by the increased clustering depicted in 2011 (**Figure 5d**). An independent 'city centre' community was not detected in any period. Instead, Manchester city centre is grouped alongside Trafford for all four censuses, and with Salford for 2011. This, we argue, is the de facto centre of the city-region, linking the commercial core, and its offshoot at Salford Quays, with the more affluent suburbs in Trafford (to the south of the city centre).

We draw two more general conclusions from our COMBO analysis. First, metropolitan communities are relatively stable, shifting only in small increments and across long time periods. Second, the modularity score decreased in each subsequent time period, suggesting that communities in Greater Manchester are becoming more interdependent.

Figure 7 shows the communities detected by InfoMap. By comparison, InfoMap detected more communities than COMBO. In 1981 and 1991, InfoMap divided Greater Manchester into nine and ten communities, respectively, which is two more than COMBO for each time period. The community structure detected is also noticeably different, with InfoMap identifying smaller communities along the northern periphery. These smaller communities sit above a larger community that grows in size between 1991 and 2011. By 2011, this large community encompasses five of the existing local authorities, exaggerating the community development that was highlighted in this area by the COMBO analysis, and again picking out the core of the Manchester conurbation.

Taking both community detection algorithms into account, the five most northern local authorities of Wigan, Bolton, Bury, Rochdale and Oldham form relatively stable communities. The lack of structural changes in these authorities, especially in the InfoMap analysis, is suggestive that they are functionally more independent than the remaining local authorities in the metropolitan area. This finding is consistent with **Folkman et al. (2016)** who argue that northern peripheral boroughs are excluded from the benefits of economic growth due to uneven development favouring the central city and inner south west quadrant.

In contrast to the stability of the northern authorities, Manchester, Salford, Stockport, Trafford and Tameside, according to the InfoMap analysis, experience significant and sustained changes in their commuting patterns and community structure. With each new year of data, these communities become more interdependent and morph into larger community structures. This analysis, then, has captured the spatial agglomeration processes that have occurred to a greater extent in the city centre and southern area of Greater Manchester, corroborating the argument of **Hodson et al. (2020)** and **Deas (2014)** that

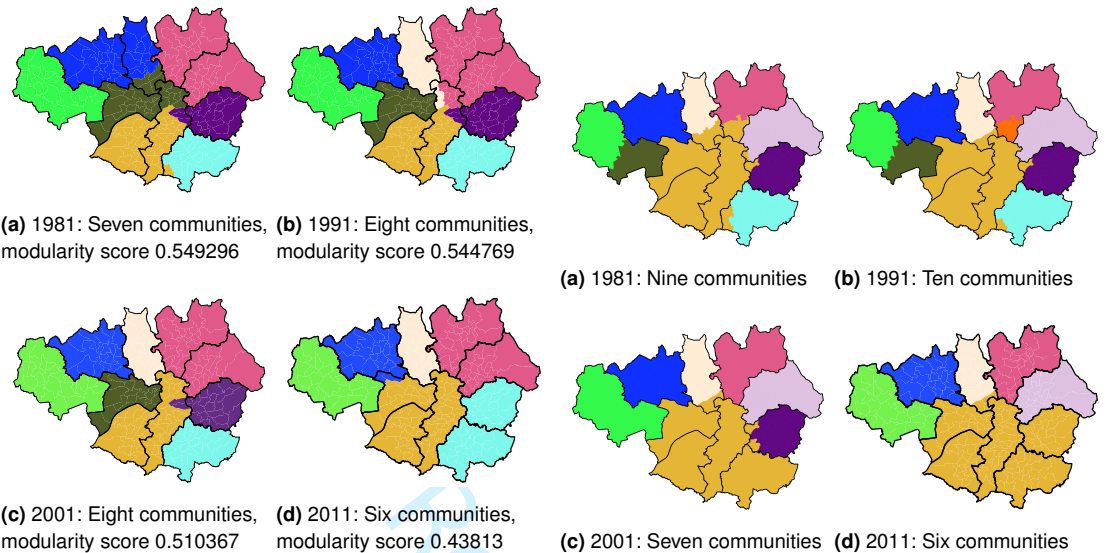


Figure 6. Comparison of the community structure detected by COMBO and the Greater Manchester local authority structure across all four periods. **Figure 7.** Comparison of the community structure detected by InfoMap and the Greater Manchester local authority structure across all four periods.

Manchester’s strategy has been to promote the geographical concentration of economic development in the core of the city-region and its southern fringes.

In the next section, we use a gravity model to refine the commuter network into a residual network. By controlling for distance and retaining a subgraph of important links, we identify the underlying economic geography of Greater Manchester.

Residual network - refining network complexity

Presenting spatial network data is challenging for sizeable (or dense) networks. This is true of Greater Manchester’s commuter network, where complex commuting patterns make initial visualisations cluttered and inhibit spatial structure from being detected. Methods to facilitate more interpretable visualisations are often based on optimising node positioning according to some centrality measure, regardless of geographic position (Hennemann 2013). For urban networks, however, maintaining a spatial element is important to understand the relationship between network topology and urban structure. To account for this, and to identify the functional economic geography of Greater Manchester, we refine the commuter network using a gravity model and plot the residual network at different thresholds.

Figure 8 shows the use of this technique in visualising the underlying structure of a spatial network. A clear image of the polycentricity of Greater Manchester emerges. At the 1% cutoff, Wigan, Bolton, and Oldham emerge as economic clusters, but the patterns across the wider region are unclear. As we increase the retained edges to 2.5% and 5%, more centres become evident. A central cluster is also depicted in

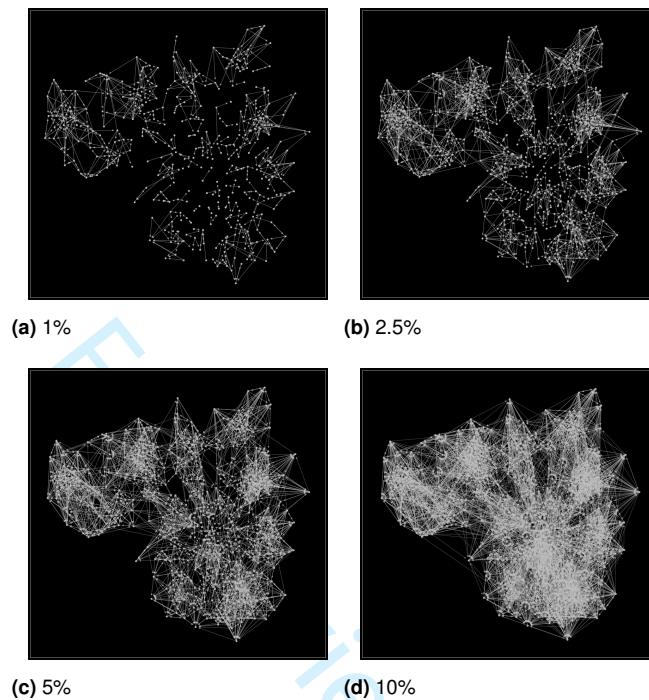


Figure 8. Residual networks for 2011 data keeping highest valued residual edges at different thresholds. Nodes are positioned geographically in respect to the centre point of their respective wards. Clusters of edges visibly reflect economic units across Greater Manchester.

both models that was captured by neither COMBO nor InfoMap. At the 10% cutoff, the visualisation becomes cluttered and makes the interpretation of functional structures increasingly difficult.

Visualising residual networks is a quick and simple method for interpreting the underlying structure of spatial networks. The residual network is essentially a subgraph of the initial commuter network that displays the most important connections across the region. Given this notion, we re-apply COMBO and InfoMap to the 5% residual network shown in Figure 8c. It is expected that the community detection methods will perform better in their respective tasks following the reduction in network complexity.

Figure 9a shows that COMBO has detected ten distinct communities, the same number as local authorities in Greater Manchester. The size, shape, and structure of the detected communities are consistent with the polycentric nature of Greater Manchester and, to a large extent, the existing local administrative boundaries. Five of the ten COMBO communities are identical to the existing administrative structure, with most other authorities seeing only slight changes. For the first time, COMBO detected a central community coterminous to Manchester city centre. This city centre community overlaps poorly with the Manchester local authority boundary, but this is unsurprising given the functionally unorthodox shape of the existing boundary. The new communities identified across the

northern border are remarkably similar to those identified in previous COMBO analysis, providing further indication that these communities are the most stable in the region.

The re-application of InfoMap to the 5% residual network produced interesting results. Figure 9b shows that whereas InfoMap previously detected an expansive community extending across central and southern Greater Manchester, we now have a number of smaller units. In refining the structure through gravity models, then, InfoMap is detecting the smaller interacting elements that form this larger community. This finding is less clear for the northern communities which are now, by comparison, larger than those in the south. A possible explanation is that the northern communities are delimited less parsimoniously at a local authority level, whereas southern communities are comprised of an amalgamation of interacting units.

When observing Figure 9a and Figure 9b together, it is difficult to ignore the role that administrative boundaries play in shaping urban structure. Despite not being in absolute agreement, all the detected communities share boundaries that align to some extent with the existing administrative borders. In the case of Wigan, for example, the COMBO analysis in Figure 9a shows that the detected community is broadly coterminous with the existing administrative boundary. InfoMap on the other hand has divided Wigan into two distinct communities. The interesting point here is that both communities remain entirely within Wigan’s administrative boundary. This happens despite the wards that are to the eastern border of Wigan being geographically as close to Salford as they are to the centre of Wigan. The question then becomes what role does the administrative boundary structure have in shaping Greater Manchester’s urban structure? Using Wigan as an example, this finding may be a reflection of local authority employment patterns, and the fact that substantial numbers of Wigan residents work for the area’s local authority. It could also be a reflection of resident perceptions of labour market opportunities, their affinity with the town and their predisposition to structure job searches to focus on localised employment. More broadly, however, questions are raised regarding how local services, planning, and funding within local authority boundaries shape the structure of metropolitan areas and perhaps fail to promote intra-regional development.

Discussion

Metropolitan areas are complex entities that are composed of multiple interacting elements. When reflecting on our results, then, it is important that we acknowledge the complexities of the system we are modelling and the limitations of our approach. The first issue in this regard is the type and temporal availability of the data used. Having only four snapshots of data over a lengthy period limits the ability to undertake sufficiently granular temporal analysis. In analysing commuter flows, we are also following previous works in suggesting that origin-destination interactions are representative of the broader patterns of connectivity that shape city-regional economic geographies (Nelson and Rae 2016; Hincks et al. 2018).

An advantage of using COMBO and InfoMap for most pieces of network-based analysis is that the algorithms determine community structure based solely on the patterns of connectivity within the network, and with no input from the user regarding the size or number of communities to be detected. In our application, however, commuter networks are dense, complex structures, with an inherent spatial underpinning. We also have an understanding of what appropriately structured administrative boundaries might look like and, as such, we can make judgements about how useful the network partitions might be in real-world debates on organising institutional and policy geometries. When reviewing the results

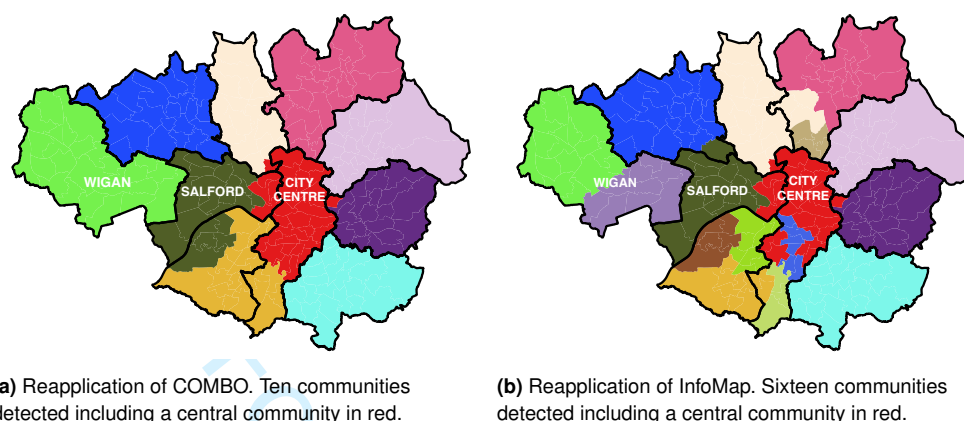


Figure 9. Comparison of the community structure detected by COMBO (a) and InfoMap (B), performed on the residual network at a 5% threshold.

presented in Figure 6 and Figure 7, then, we see that COMBO and InfoMap have achieved, to some extent, promising results in identifying spatial units that are coterminous to the existing administrative boundary structure. However, following the use of gravity models to refine the complexity of the network, Figure 9a shows that COMBO has now captured a more fragmented community structure that is closer to known clustering phenomena of different socioeconomic drivers. Using gravity models to create residual networks prior to performing community detection methods has improved results in our application. In this way, we liken the use of gravity models to that of feature engineering applications in machine learning, where domain knowledge is leveraged to improve the performance of computational tasks.

Whilst it is evident in our analysis that administrative boundaries are important structures in defining intra-regional economic geographies, as our results are comparable with existent administrative boundaries, the question remains as to why this happens to the extent we have seen. In instances where administrative boundaries do not align well with the geography of detected communities, the difficulty of reorganising boundaries must not be overlooked. Such undertakings are eclipsed by more pressing policy reform issues that are more easily delivered. This argument stands not only from a regional perspective, where the focus is on what is in a region, but also from an intra-regional perspective where the question of who is responsible for delivering local services becomes the primary concern. Against the backdrop of austerity and through work consolidated by AGMA, some of the Greater Manchester authorities have previously considered joint-working as a way of achieving economies of scale in relation to local public service delivery. The intra-regional geographies identified in this work might suggest particular alliances that make sense in relation to joint-working on specific sectors and policies.

Political, social and cultural factors will ultimately define the composition of local authorities, but this work has demonstrated the value of a network approach in providing evidence-based support to inform the development of intra-regional institutional structures and policies.

Conclusions

The findings from this research suggest that a combination of network analysis, spatial interaction modelling, and geographical network visualisation can identify the underlying urban morphology of metropolitan regions. We argue that without network refinement using gravity models, community detection methods are challenged with unpicking the complex interdependencies that are a characteristic feature of intra-urban networks. The empirical analysis of the Greater Manchester region revealed a community structure that was geographically and algorithmically coherent and has the potential to be used as a framework to inform a policy debate over local service delivery. We have also demonstrated the importance of administrative boundaries in shaping the economic geography of administrative regions.

This work suggests two main directions for future research. First, this framework of analysis could be applied to different metropolitan settings, geographic scales, and would benefit from higher resolution spatial and temporal data that is likely to become more readily available. Second, commuter networks stand as an important proxy for intra-regional flows, but they say little about other relationships that shape urban areas (for example, professional, social or cultural networks). Extending beyond a monoplex (i.e. single-layer) network framework, could allow interrelations with urban other networks to be explored. Emerging work in the field of network science is being used to model systems of multiple interacting elements using multilayer networks (Bianconi 2018). Whilst there are complexities in defining meaningful relationships between networks of different nodes, the potential to understand how urban networks co-evolve and influence each other is an area of potential future research.

References

Barthélemy M (2011) Spatial networks. *Physics Reports* 499(1-3): 1–101.

Batty M (2013) *The new science of cities*. MIT press.

Berdegúe JA, Hiller T, Ramírez JM, Satizábal S, Soloaga I, Soto J, Uribe M and Vargas O (2019) Delineating functional territories from outer space. *Latin American Economic Review* 28(1): 4.

Bianconi G (2018) *Multilayer networks: structure and function*. Oxford university press.

Boeing G (2019) Urban spatial order: Street network orientation, configuration, and entropy. *Applied Network Science* 4(1): 67.

Cheng YY, Lee RKW, Lim EP and Zhu F (2015) Measuring centralities for transportation networks beyond structures. In: *Applications of social media and social network analysis*. Springer, pp. 23–39.

Csardi G and Nepusz T (2006) The igraph software package for complex network research. *InterJournal Complex Systems*: 1695. URL <https://igraph.org>.

Deas I (2014) The search for territorial fixes in subnational governance: City-regions and the disputed emergence of post-political consensus in manchester, england. *Urban Studies* 51(11): 2285–2314.

Demazière C and Sykes O (2020) The rise of the metropolitan city region? exploring the establishment of new levels of local government in england and france. In: *Foregrounding Urban Agendas*. Springer, pp. 185–209.

Esquivel AV and Rosvall M (2011) Compression of flow can reveal overlapping-module organization in networks. *Physical Review X* 1(2): 021025.

Fang C and Yu D (2017) Urban agglomeration: An evolving concept of an emerging phenomenon. *Landscape and Urban Planning* 162: 126–136.

- Folkman P, Froud J, Johal S, Tomaney J and Williams K (2016) Manchester transformed: why we need a reset of city region policy. *CRESC Public Interest Report, November*.
- Hagberg A, Swart P and S Chult D (2008) Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States).
- Hamilton R and Rae A (2018) Regions from the ground up: a network partitioning approach to regional delineation. *Environment and Planning B: Urban Analytics and City Science* : 2399808318804226.
- Harding A, Harloe M and Rees J (2010) Manchester's bust regime? *International Journal of Urban and Regional Research* 34(4): 981–991.
- Hartshorne R (1939) The nature of geography: a critical survey of current thought in the light of the past (conclusion). *Annals of the Association of American Geographers* 29(4): 413–658.
- Haughton G, Deas I, Hincks S and Ward K (2016) Mythic manchester: Devo manc, the northern powerhouse and rebalancing the english economy. *Cambridge Journal of Regions, Economy and Society* 9(2): 355–370.
- He M, Glasser J, Pritchard N, Bhamidi S and Kaza N (2019) Demarcating geographic regions using community detection in commuting networks. *arXiv preprint arXiv:1903.06029*.
- Hennemann S (2013) Information-rich visualisation of dense geographical networks. *Journal of maps* 9(1): 68–75.
- Hincks S (2012) Daily interaction of housing and labour markets in north west england. *Regional Studies* 46(1): 83–104.
- Hincks S, Kingston R, Webb B and Wong C (2018) A new geodemographic classification of commuting flows for england and wales. *International Journal of Geographical Information Science* 32(4): 663–684.
- Hodson M, McMeekin A, Froud J and Moran M (2020) State-rescaling and re-designing the material city-region: Tensions of disruption and continuity in articulating the future of greater manchester. *Urban Studies* 57(1): 198–217.
- Jiang B (2018) Geospatial analysis requires a different way of thinking: The problem of spatial heterogeneity. In: *Trends in Spatial Analysis and Modelling*. Springer, pp. 23–40.
- Jiang B and Claramunt C (2004) Topological analysis of urban street networks. *Environment and Planning B: Planning and design* 31(1): 151–162.
- Kenealy D (2016) A tale of one city: The devo manc deal and its implications for english devolution. *The Political Quarterly* 87(4): 572–581.
- Louail T, Lenormand M, Ros OGC, Picornell M, Herranz R, Frias-Martinez E, Ramasco JJ and Barthelemy M (2014) From mobile phone data to the spatial structure of cities. *Scientific reports* 4: 5276.
- Nelson GD and Rae A (2016) An economic geography of the united states: From commutes to megaregions. *PloS one* 11(11): e0166083.
- Newman M (2010) *Networks: An Introduction*. Oxford: Oxford University Press. ISBN 9780199206650. DOI: 10.1093/acprof:oso/9780199206650.001.0001.
- Pflieger G and Rozenblat C (2010) Introduction. urban networks and network theory: the city as the connector of multiple networks.
- Plane DA (1981) The geography of urban commuting fields: some empirical evidence from new england. *The Professional Geographer* 33(2): 182–188.
- Porta S, Crucitti P and Latora V (2006) The network analysis of urban streets: a dual approach. *Physica A: Statistical Mechanics and its Applications* 369(2): 853–866.
- Rodrigue JP (2016) *The geography of transport systems*. Taylor & Francis.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

Detecting shifts in metropolitan structure

17

Rosvall M, Axelsson D and Bergstrom CT (2009) The map equation. *The European Physical Journal Special Topics* 178(1): 13–23.

Roth C, Kang SM, Batty M and Barthélemy M (2011) Structure of urban movements: polycentric activity and entangled hierarchical flows. *PloS one* 6(1): e15923.

Sng TH, Chia PZ, Feng CC and Wang YC (2018) Are china’s provincial boundaries misaligned? *Applied Geography* 98: 52–65.

Sobolevsky S, Campari R, Belyi A and Ratti C (2014) General optimization technique for high-quality community detection in complex networks. *Physical Review E* 90(1): 012811.

Stillwell J (2006) Providing access to census-based interaction data in the uk: that’s wicid. *The Journal of Systemics, Cybernetics and Informatics* 4(4): 63–68.

Thiemann C, Theis F, Grady D, Brune R and Brockmann D (2010) The structure of borders in a small world. *PloS one* 5(11): e15422.

Zhong C, Arisona SM, Huang X, Batty M and Schmitt G (2014) Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science* 28(11): 2178–2199.

Supplementary material

Methods

Average degree. The average degree, k , of a directed network is calculated as

$$k = \frac{m}{n}$$

where m and n are the total edges and total nodes, respectively.

Network density. Network density describes the prevalence of connections within a network and is a useful measure of analysing network structure. It is calculated by taking the actual edges within a network as a portion of the potential edges, formally given as

$$D = \frac{m}{n(n-1)}.$$

In the remainder of this section, we present the centrality measures and algorithms used in our analysis. They are: node strength, eigenvector centrality, PageRank, HITS and clustering coefficient.

Node strength. The weighted degree, also called node strength, of a node in a network measures the sum of weighted edges incident on a node. For directed networks, we define the weighted-in- degree and weighted-out-degree as

$$s_i^{in} = \sum_j^N w_{ij},$$

$$s_i^{out} = \sum_j^N w_{ji}.$$

where w_{ij} and w_{ji} are the weights of connections that point to and from node i , respectively. These weighted degree measures only consider the strength of connections incident on a node but do not account for the relative importance of such connections or their qualitative substance.

Eigenvector centrality. The eigenvector centrality is a measure that quantifies a node's importance in a network by considering its connections to other nodes that are themselves important (Newman, 2010). It does this by giving each node a score proportional to the sum of scores of its neighbours. Thus, if the importance of the N nodes of the network is quantified by the vector $x \in \mathbb{R}^N$, $x = (x_1, \dots, x_i, \dots, x_N)^T$ then the importance of node i is defined by:

$$x_i = \frac{1}{\lambda} \sum_{j=1}^N a_{ij} x_j,$$

where $\lambda \in \mathbb{R}$ λ is a constant such that $\lambda \neq 0$, and a_{ji} are the elements of matrix $A \in \mathbb{R}^{N \times N}$, with A the adjacency matrix. For all the nodes in the network, we have:

$$\lambda x = xA.$$

PageRank. PageRank accounts for edge direction. PageRank uses the structure of incoming links to rank nodes, with a node considered important if it is linked from other important nodes or if it is highly linked. Following (Newman, 2010), the PageRank centrality x_i of node i is given by

$$x_i = \alpha \sum_k \frac{\alpha_{ki}}{d_k} x_k + B$$

where α and B are constants and d_k is the out-degree of node k if such degree is positive, or $d_k = 1$ if the out-degree of k is null.

HITS –hub and authority scores. The HITS (hypertext induced topic selection) algorithm is similar to PageRank in that it was originally designed to rank web pages based on their directed wiring. In this regard, HITS assigns two scores for each node: its authority x_i which estimates the value of the node and is most relevant in the study of commuter networks, and its hub value y_i which estimates the value of its links to other nodes. The two scores are intrinsically linked: a node has a high authority value if it is linked to by many nodes with a high hub value, and a node has a high hub value if it links to many nodes with a high authority value. Using this notion, (Kleinberg, 1999) formally defines the authority score and hub score as

$$x_i = \sum_{j=1}^N y_j,$$

$$y_i = \sum_{j=1}^N x_j.$$

Network density and clustering coefficient. The clustering coefficient is a measure of the degree to which wards in the commuter network tend to cluster together. We expect to see tightly knit groups form around economic centres. For directed and weighted networks, (Fagiolo, 2007) formally defines the clustering coefficient as:

$$c_u = \frac{1}{\deg^{tot}(u)(\deg^{tot}(u) - 1) - 2\deg^{\leftrightarrow}(u)} T(u)$$

where T_u is the number of directed triangles through node u , $\deg^{\{tot\}}(u)$ is the sum of in degree and out degree of u and $\deg^{\leftrightarrow}(u)$ is the reciprocal degree of u , defined as the ratio of the number of edges in both directions to the total number of edges attached to node u .

Bibliography

- Fagiolo, G., 2007. Clustering in complex directed networks. *Physical review E*, 76(2).
 Kleinberg, J. M., 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5), pp. 604-632.
 Newman, M., 2010. *Networks: An Introduction*. Oxford: Oxford University Press.